

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/479,467DATE: 02/01/2000  
TIME: 11:29:16

Input Set: I479467.RAW

This Raw Listing contains the General Information Section and up to first 5 pages.

ENTERED

1 <110> APPLICANT: California Institute of Technology et al.  
 2 <120> TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE  
 3 BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON  
 4 <130> FILE REFERENCE: 18021-2901PC  
 5 <140> CURRENT APPLICATION NUMBER: US/09/479,467  
 6 <141> CURRENT FILING DATE: 2000-01-06  
 7 <150> EARLIER APPLICATION NUMBER: 60/115,127  
 8 <151> EARLIER FILING DATE: 1999-01-06  
 9 <160> NUMBER OF SEQ ID NOS: 16  
 10 <170> SOFTWARE: PatentIn Ver. 2.0  
 11 <210> SEQ ID NO 1  
 12 <211> LENGTH: 12912  
 13 <212> TYPE: DNA  
 14 <213> ORGANISM: Homo sapiens PKD-1 gene  
 15 <220> FEATURE:  
 16 <221> NAME/KEY: CDS  
 17 <222> LOCATION: (1)..(12912)  
 18 <300> PUBLICATION INFORMATION:  
 19 <310> PATENT DOCUMENT NUMBER: 5,891,628  
 20 <311> PATENT FILING DATE: 1995-06-02  
 21 <312> PUBLICATION DATE: 1999-04-06  
 22 <400> SEQUENCE: 1 48  
 atg ccg ccc gcc ggc ccc gcc cgcc ctg ggc ctg ggc ctg ggc  
 23 Met Pro Pro Ala Ala Pro Ala Arg Leu Ala Leu Ala Leu Gly Leu Gly 48  
 24 Met Pro Pro Ala Ala Pro Ala Arg Leu Ala Leu Ala Leu Gly Leu Gly 15  
 25 1 5 10 96  
 26 ctg tgg ctc ggg ggc ctg ggc ggg ggg ggg ccc ggc ggc tgc ggg ccc 96  
 27 Leu Trp Leu Gly Ala Leu Ala Gly Gly Pro Gly Arg Gly Cys Gly Pro 30  
 28 20 25 30 144  
 29 tgc gag ccc ccc tgc ctc tgc ggg cca ggc ccc ggc gcc tgc cgc 144  
 30 Cys Glu Pro Pro Cys Leu Cys Gly Pro Ala Pro Gly Ala Ala Cys Arg 45  
 31 35 40 45 192  
 32 gtc aac tgc tcg ggc cgc ggg ctg cgg acg ctc ggt ccc ggc ctg cgc 192  
 33 Val Asn Cys Ser Gly Arg Gly Leu Arg Thr Leu Gly Pro Ala Leu Arg 50 55 60  
 34 atc ccc ggc gac aca gag cta gac gtc tcc cac aac ctg ctc cgg 240  
 35 Ile Pro Ala Asp Ala Thr Glu Leu Asp Val Ser His Asn Leu Leu Arg 65 70 75 80  
 36 80 288  
 37 gcg ctg gac gtt ggg ctc ctg ggc aac ctc tgc ggc ctg gca gag ctg 288  
 38 Ala Leu Asp Val Gly Leu Leu Ala Asn Leu Ser Ala Leu Ala Glu Leu 85 90 95  
 39 100 105 110 336  
 40 gat ata agc aac aac aag att tct acg tta gaa gaa ata ttt gct 336  
 41 Asp Ile Ser Asn Asn Lys Ile Ser Thr Leu Glu Glu Gly Ile Phe Ala 100 105 110 384  
 42 43 aat tta ttt aat tta agt gaa ata aac ctg agt ggg aac ccg ttt gag 44

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/479,467DATE: 02/01/2000  
TIME: 11:29:16

Input Set: I479467.RAW

45	Asn Leu Phe Asn Leu Ser Glu Ile Asn Leu Ser Gly Asn Pro Phe Glu				
46	115	120	125		
47	tgt gac tgt ggc ctc gcg tgg ctg cgg caa tgg gcg gag gag cag cag			432	
48	Cys Asp Cys Gly Leu Ala Trp Leu Pro Gln Trp Ala Glu Glu Gln Gln				
49	130	135	140		
50	gtg cgg tgg gtc cag ccc gag gca gcc acg tgg gct ggg cct ggc tcc			480	
51	Val Arg Val Val Glu Pro Glu Ala Ala Thr Cys Ala Gly Pro Gly Ser				
52	145	150	155	160	528
53	ctg gct ggc cag cct ctg ctt ggc atc ccc ttg ctg gac agt ggc tgg				
54	Leu Ala Gly Gln Pro Leu Leu Gly Ile Pro Leu Leu Asp Ser Gly Cys				
55	165	170	175		
56	ggt gag gag tat gtc gcc tgc ctc oct gac aac agc tca ggc acc gtc			576	
57	Gly Glu Tyr Val Ala Cys Leu Pro Asp Asn Ser Ser Gly Thr Val				
58	180	185	190		
59	gca gca gtc tcc ttt tca gct gcc cac gaa ggc ctc ctt cag cca gag			624	
60	Ala Ala Val Ser Phe Ser Ala Ala His Glu Gly Leu Leu Gln Pro Glu				
61	195	200	205		
62	gcc tgc agc gcc ttc tgc ttc acc ggc cag ggc ctc gca gcc ctc			672	
63	Ala Cys Ser Ala Phe Cys Ser Thr Gly Glu Gly Leu Ala Ala Leu				
64	210	215	220		
65	tcg gag cag ggc tgg tgc ctg tgt ggg gcg cag ccc tcc agt gcc			720	
66	Ser Glu Gln Gly Trp Cys Leu Cys Gly Ala Ala Gln Pro Ser Ser Ala				
67	225	230	235	240	768
68	tcc ttt gcc tgc ctg tcc ctc tgc tcc ggg ccc ccc gca cct cct gcc				
69	Ser Phe Ala Cys Leu Ser Leu Cys Ser Gly Pro Pro Ala Pro Pro Ala				
70	245	250	255		
71	ccc acc tgt agg ggc ccc acc ctc ctc cag cac gtc ttc cct gcc tcc			816	
72	Pro Thr Cys Arg Gly Pro Thr Leu Leu Gln His Val Phe Pro Ala Ser				
73	260	265	270		
74	cca ggg gcc acc ctg gtc ggg ccc cac gga cct ctg gcc tct ggc cag			864	
75	Pro Gly Ala Thr Leu Val Gly Pro His Gly Pro Leu Ala Ser Gly Gln				
76	275	280	285		
77	cta gca gcc ttc cac atc gct gcc ccc ctc cct gtc act gac aca cgc			912	
78	Leu Ala Ala Phe His Ile Ala Ala Pro Leu Pro Val Thr Asp Thr Arg				
79	290	295	300		
80	tgg gac ttc gga gac ggc tcc gcc gag gtc gat gcc gct ggg ccc gct				
81	Trp Asp Phe Gly Asp Gly Ser Ala Glu Val Asp Ala Ala Gly Pro Ala				
82	305	310	315	320	
83	gcc tcg cat cgc tat gtc ctg cct ggg cgc tat cac gtc acg gcc gtc				
84	Ala Ser His Arg Tyr Val Leu Pro Gly Arg Tyr His Val Thr Ala Val				
85	325	330	335		
86	ctg gcc ctg ggg ggc ggc tca gcc ctg ctg ggg aca gac gtc cag gtc			1056	
87	Leu Ala Leu Gly Ala Gly Ser Ala Leu Leu Gly Thr Asp Val Gln Val				
88	340	345	350		
89	gaa gcg gca cct gcc gcc ctg gag ctc gtc tgg ccg tcc tcg gtc gag			1104	
90	Glu Ala Ala Pro Ala Ala Leu Glu Leu Val Cys Pro Ser Ser Val Gln				
91	355	360	365		
92	agt gac gag agc ctc gac ctc agc atc cag aac cgc ggt ggt tca ggc			1152	
93	Ser Asp Glu Ser Leu Asp Leu Ser Ile Gln Asn Arg Gly Gly Ser Gly				
94	370	375	380		

Input Set: I479467.RAW

95 ctg gag gcc gcc tac agc atc gtg gcc ctg ggc gag gag cgg gcc cga 1200  
 96 Leu Glu Ala Ala Tyr Ser Ile Val Ala Leu Gly Glu Glu Pro Ala Arg  
 97 385 390 395 400  
 98 gcg gtg cac ccg ctc tgc ccc tcg gac acg gag atc ttc cct ggc aac 1248  
 99 Ala Val His Pro Leu Cys Pro Ser Asp Thr Glu Ile Phe Pro Gly Asn  
 100 405 410 415  
 101 ggg cac tgc tac cgc ctg gtg gag aag ggc gcc tgg ctg cag cgc 1296  
 102 Gly His Cys Tyr Arg Leu Val Val Glu Lys Ala Ala Trp Leu Gln Ala  
 103 420 425 430  
 104 cag gag cag tgt cag gcc tgg gcc ggg gcc ctc gca atg gtg gac 1344  
 105 Gln Glu Gln Cys Gln Ala Trp Ala Gly Ala Ala Leu Ala Met Val Asp  
 106 435 440 445  
 107 agt ccc gcc gtg cag cgc ttc ctg gtc tcc cgg gtc acc agg agc cta 1392  
 108 Ser Pro Ala Val Gln Arg Phe Leu Val Ser Arg Val Thr Arg Ser Leu  
 109 450 455 460  
 110 gac gtg tgg atc ggc ttc tgc act gtg cag ggg gtg gag gtc ggc cca 1440  
 111 Asp Val Trp Ile Gly Phe Ser Thr Val Gln Gly Val Glu Val Gly Pro  
 112 465 470 475 480  
 113 gcg ccc cag ggc gag gcc ttc agc ctg gag agc tgc cag aac tgg ctg 1488  
 114 Ala Pro Gln Gly Glu Ala Phe Ser Leu Glu Ser Cys Gln Asn Trp Leu  
 115 485 490 495  
 116 ccc ggg gag cca cac cca gcc aca gcc gag cac tgc gtc cgg ctc ggg 1536  
 117 Pro Gly Glu Pro His Pro Ala Thr Ala Glu His Cys Val Arg Leu Gly  
 118 500 505 510  
 119 ccc acc ggg tgg tgt aac acc gac ctg tgc tca gcg ccc cac agc tac 1584  
 120 Pro Thr Gly Trp Cys Asn Thr Asp Leu Cys Ser Ala Pro His Ser Tyr  
 121 515 520 525  
 122 gtc tgc gag cag ccc gga cca gtg cag gat gcc gag aac ctc 1632  
 123 Val Cys Leu Gln Pro Gly Pro Val Gln Asp Ala Glu Asn Leu  
 124 530 535 540  
 125 ctc gtg gga gcg ccc agt ggg gac ctg cag gga ccc ctg acg cct ctg 1680  
 126 Leu Val Gly Ala Pro Ser Gly Asp Leu Gln Gly Pro Leu Thr Pro Leu  
 127 545 550 555 560  
 128 gca cag cag gag ggc ctc tca gcc ccc cac gag ccc gtg gag gtc atg 1728  
 129 Ala Gln Gln Asp Gly Leu Ser Ala Pro His Glu Pro Val Glu Val Met  
 130 565 570 575  
 131 gta ttc ccg ggc ctg cgt ctg agc cgt gaa gcc ttc ctc acc acg gcc 1776  
 132 Val Phe Pro Gly Leu Arg Leu Ser Arg Glu Ala Phe Leu Thr Thr Ala  
 133 580 585 590  
 134 gaa ttt ggg acc cag gag ctc cgg cgg ccc gcc cag ctg cgg ctg cag 1824  
 135 Glu Phe Gly Thr Gln Glu Leu Arg Arg Pro Ala Gln Leu Arg Leu Gln  
 136 595 600 605  
 137 gtg tac cgg ctc ctc agc aca gca ggg acc ccc gag aac ggc agc gag 1872  
 138 Val Tyr Arg Leu Leu Ser Thr Ala Gly Thr Pro Glu Asn Gly Ser Glu  
 139 610 615 620  
 140 cct gag agc agg tcc ccg gac aac agg acc cag ctg gcc ccc ggc tgc 1920  
 141 Pro Glu Ser Arg Ser Pro Asp Asn Arg Thr Gln Leu Ala Pro Ala Cys  
 142 625 630 635 640  
 143 atg cca ggg gga cgc tgg tgc cct gga gcc aac atc tgc ttg cgg ctg 1968  
 144 Met Pro Gly Gly Arg Trp Cys Pro Gly Ala Asn Ile Cys Leu Pro Leu

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/479,467DATE: 02/01/2000  
TIME: 11:29:16

Input Set: I479467.RAW

145	645	650	655	2016
146	gac gcc tcc tgc cac ccc cag gcc tgc ggc aat ggc tgc acg tca ggg			
147	Asp Ala Ser Cys His Pro Gln Ala Cys Ala Asn Gly Cys Thr Ser Gly			
148	660	665	670	2064
149	cca ggg cta ccc ggg gcc ccc tat gcg cta tgg aga gag ttc ctc ttc			
150	Pro Gly Leu Pro Gly Ala Pro Tyr Ala Leu Trp Arg Glu Phe Leu Phe			
151	675	680	685	2112
152	tcc gtt ccc gcg ggg ccc ccc gcg cag tac tcg gtc acc ctc cac ggc			
153	Ser Val Pro Ala Gly Pro Ala Gln Tyr Ser Val Thr Leu His Gly			
154	690	695	700	2160
155	cag gat gtc ctc atg ctc cct ggt gac ctc gtt ggc ttg cag cac gac			
156	Gln Asp Val Leu Met Leu Pro Gly Asp Leu Val Gly Leu Gln His Asp			
157	705	710	715	2208
158	gct ggc cct ggc gcc ctc ctg cac tgc tgc ccg gtc ccc ggc cac cct			
159	Ala Gly Pro Gly Ala Leu Leu His Cys Ser Pro Ala Pro Gly His Pro			
160	725	730	735	2256
161	ggt ccc cgg gcc ccg tac ctc tcc gcc aac gcc tgc tca tgg ctg ccc			
162	Gly Pro Arg Ala Pro Tyr Leu Ser Ala Asn Ala Ser Ser Trp Leu Pro			
163	740	745	750	2304
164	cac ttg cca gcc cag ctg gag ggc act tgg ggc tgc cct gcc tgt gcc			
165	His Leu Pro Ala Gln Leu Glu Gly Thr Trp Gly Cys Pro Ala Cys Ala			
166	755	760	765	2352
167	ctg cgg ctg ctt gca caa ccg gaa cag ctc acc gtg ctg ctg ggc ttg			
168	Leu Arg Leu Leu Ala Gln Arg Glu Gln Leu Thr Val Leu Leu Gly Leu			
169	770	775	780	2400
170	agg ccc aac cct gga ctg cgg ctg cct ggg cgc tat gag gtc cgg gca			
171	Arg Pro Asn Pro Gly Leu Arg Leu Pro Gly Arg Tyr Glu Val Arg Ala			
172	785	790	795	2448
173	gag gtg ggc aat ggc gtg tcc agg cac aac ctc tcc tgc agc ttt gac			
174	Glu Val Gly Asn Gly Val Ser Arg His Asn Leu Ser Cys Ser Phe Asp			
175	805	810	815	2496
176	gtg gtc tcc cca gtg gct ggg ctg cgg gtc atc tac ctc cct gcc ccc cgc			
177	Val Val Ser Pro Val Ala Gly Leu Arg Val Ile Tyr Pro Ala Pro Arg			
178	820	825	830	2544
179	gac ggc cgc ctc tac gtg ccc acc aac ggc tca gcc ttg gtg ctc cag			
180	Gly Arg Leu Tyr Val Pro Thr Asn Gly Ser Ala Leu Val Leu Gln			
181	835	840	845	2592
182	gtg gac tct ggt gcc aac gcc acg gcc acg gct cgc tgg cct ggg ggc			
183	Val Asp Ser Gly Ala Asn Ala Thr Ala Thr Ala Arg Trp Pro Gly Gly			
184	850	855	860	2640
185	agt ctc agc gcc cgc ttt gag aat gtc cct gcc ctg gtg gcc acc			
186	Ser Leu Ser Ala Arg Phe Glu Asn Val Cys Pro Ala Leu Val Ala Thr			
187	865	870	875	2688
188	ttc gtg ccc gcc tgc ccc tgg gag acc aac gat acc ctg ttc tca gtg			
189	Phe Val Pro Ala Cys Pro Trp Glu Thr Asn Asp Thr Leu Phe Ser Val			
190	885	890	895	2736
191	gta gca ctg cgg tgg ctc agt gag ggg gag cac gtg gtg gac gtg gtg			
192	Val Ala Leu Pro Trp Leu Ser Glu Gly Glu His Val Val Asp Val Val			
193	900	905	910	2784
194	gtg gaa aac agc gcc agc cgg gcc aac ctc agc ctg cgg gtg acg gcg			

PAGE: 5

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/479,467DATE: 02/01/2000  
TIME: 11:29:16

Input Set: I479467.RAW

195	Val	Glu	Asn	Ser	Ala	Ser	Arg	Ala	Asn	Leu	Ser	Leu	Arg	Val	Thr	Ala		
196		915								920			925					
197	gag	gag	ccc	atc	tgt	ggc	ctc	cgc	gcc	acg	ccc	agg	ccc	gag	gcc	cgt	2832	
198	Glu	Glu	Pro	Ile	Cys	Gly	Leu	Arg	Ala	Thr	Pro	Ser	Pro	Glu	Ala	Arg		
199		930								935			940					
200	gta	ctg	tag	ggg	gtc	cta	gtg	agg	tac	agc	ccc	gtg	gtg	gag	gcc	ggc	2880	
201	Val	Leu	Gln	Gly	Val	Leu	Val	Arg	Tyr	Ser	Pro	Val	Val	Glu	Ala	Gly		
202		945								950			955					
203	tcg	gac	atg	gtc	tcc	cgg	tgg	acc	atc	aac	gac	aag	cag	tcc	ctg	acc	2928	
204	Ser	Asp	Met	Val	Phe	Arg	Trp	Thr	Ile	Asn	Asp	Lys	Gln	Ser	Leu	Thr		
205		965								970			975					
206	tcc	cag	gtg	gtc	tcc	aat	gtc	att	tat	cag	agc	gcg	gct	gtc	tcc		2976	
207	Phe	Gln	Asn	Val	Val	Phe	Asn	Val	Ile	Tyr	Gln	Ser	Ala	Ala	Val	Phe		
208		980								985			990					
209	aag	ctc	tca	ctg	acg	gcc	tcc	aac	cac	gtg	agc	aac	gtc	acc	gtg	aac	3024	
210	Lys	Leu	Ser	Leu	Thr	Ala	Ser	Asn	His	Val	Ser	Asn	Val	Thr	Val	Asn		
211		995								1000			1005					
212	tac	aac	gtt	gtg	gag	cg	atg	aac	agg	atg	cag	qgt	ctg	cag	gtc		3072	
213	Tyr	Asn	Val	Thr	Val	Glu	Arg	Met	Asn	Arg	Met	Gln	Gly	Leu	Gln	Val		
214		1010								1015			1020					
215	tcc	aca	gtt	gtg	ccc	gtt	tcc	aat	gcc	acg	cta	gca	ctg	acg			3120	
216	Ser	Thr	Val	Pro	Ala	Val	Leu	Ser	Pro	Asn	Ala	Thr	Leu	Ala	Leu	Thr		
217		1025								1030			1035					3168
218	gct	ggc	gtt	gtg	gtt	gac	tcc	gcc	gtg	gag	gtg	gcc	tcc	ctg	tgg	acc		
219	Ala	Gly	Val	Leu	Val	Asp	Ser	Ala	Val	Ala	Phe	Leu	Trp	Thr				
220		1045								1050			1055					
221	ttt	ggg	gat	ggg	gag	cag	gcc	ctc	cac	cag	ttc	cag	ccg	ctg	cac		3216	
222	Phe	Gly	Asp	Gly	Glu	Gln	Ala	Leu	His	Gln	Phe	Gln	Pro	Pro	Tyr	Asn		
223		1060								1065			1070					
224	gag	tcc	tcc	cca	gtt	cca	gac	ccc	tcc	gtg	gcc	cag	gtg	ctg	gtg	gag	3264	
225	Glu	Ser	Pro	Val	Pro	Val	Pro	Asp	Pro	Ser	Val	Ala	Val	Leu	Val	Glu		
226		1075								1080			1085					
227	cac	aat	gtc	acg	cac	acc	tac	gtt	gcc	cca	gtt	gag	tac	ctc	ctg	acc	3312	
228	His	Asn	Val	Thr	His	Thr	Tyr	Ala	Ala	Pro	Gly	Glu	Tyr	Leu	Leu	Thr		
229		1090								1095			1100					
230	gtg	ctg	gca	tct	aat	gcc	ttc	gag	aac	ctg	acg	cag	gtg	cct	gtg		3360	
231	Val	Leu	Ala	Ser	Asn	Ala	Phe	Glu	Asn	Leu	Thr	Gln	Gln	Val	Pro	Val		
232		1105								1110			1115					3408
233	agc	gtt	gtc	ccg	gcc	tcc	ccc	tcc	gtt	gtt	gtt	gtt	gtt	gtt	gtt	gtt		
234	Ser	Val	Arg	Ala	Ser	Leu	Pro	Ser	Val	Ala	Val	Gly	Val	Ser	Asp	Gly		
235		1125								1130			1135					3456
236	gtc	ctg	gtt	gcc	ggc	ccg	ccc	gtc	acc	ttc	tac	ccg	cac	ccg	ctg	ccc		
237	Val	Leu	Val	Ala	Gly	Arg	Pro	Val	Thr	Phe	Tyr	Pro	His	Pro	Leu	Pro		
238		1140								1145			1150					3504
239	tcg	cct	ggg	ggg	gtt	ttt	tac	acg	tgg	gac	ttc	ggg	gac	ggc	tcc	cct		
240	Ser	Pro	Gly	Gly	Val	Leu	Tyr	Thr	Trp	Asp	Phe	Gly	Asp	Gly	Ser	Pro		
241		1155								1160			1165					3552
242	gtc	ctg	acc	cag	acg	cag	ccg	gtt	gcc	aac	cac	acc	tat	gcc	tcg	agg		
243	Val	Leu	Thr	Gln	Ser	Gln	Pro	Ala	Ala	Asn	His	Thr	Tyr	Ala	Ser	Arg		
244		1170								1175			1180					

PAGE: 6

VERIFICATION SUMMARY  
PATENT APPLICATION US/09/479,467

DATE: 02/01/2000  
TIME: 11:29:16

Input Set: I479467.RAW

Line ? Error/Warning

Original Text